D684–D688 *Nucleic Acids Research, 2006, Vol. 34, Database issue*
doi:10.1093/nar/gkj155

RTPrimerDB: the real-time PCR primer and probe database, major update 2006

Filip Pattyn, Piet Robbrecht, Anne De Paepe, Frank Speleman and Jo Vandesompele*

Center for Medical Genetics Ghent (CMGG), Ghent University Hospital, De Pintelaan 185, 9000 Ghent, Belgium

Received September 15, 2005; Revised and Accepted October 31, 2005

ABSTRACT

The RTPrimerDB (http://medgen.ugent.be/rtprimerdb) project provides a freely accessible data retrieval system and an *in silico* assay evaluation pipeline for real-time quantitative PCR assays. Over the last year the number of user submitted assays has grown to 3500. Data conveyance from Entrez Gene by establishing an assay-to-gene relationship enables the addition of new primer assays for one of the 1.5 million different genes from 2300 species stored in the system. Easy access to the primer and probe data is possible by using multiple search criteria. Assay reports contain gene information, assay details (such as oligonucleotide sequences, detection chemistry and reaction conditions), publication information, users' experimental evaluation feedback and submitter's contact details. Gene expression assays are extended with a scalable assay viewer that provides detailed information on the alignment of primer and probe sequences on the known transcript variants of a gene, along with Single Nucleotide Polymorphisms (SNP) positions and peptide domain information. Furthermore, an mfold module is implemented to predict the secondary structure of the amplicon sequence, as this has been reported to impact the efficiency of the PCR. RTPrimerDB is also extended with an *in silico* analysis pipeline to streamline the evaluation of custom designed primer and probe sequences prior to ordering and experimental evaluation. In a secured environment, the pipeline performs automated BLAST specificity searches, mfold secondary structure prediction, SNP or plain sequence error identification, and graphical visualization of the aligned primer and probe sequences on the target gene.

INTRODUCTION

RTPrimerDB is a public primer and probe database for real-time quantitative PCR (qPCR) applications hosted at the Center for Medical Genetics in the Ghent University Hospital in Gent, Belgium. The database was initially developed to address the problem of the repeated laborious primer design and assay evaluation for quantification or detection of the same nucleic acid sequences by different individuals, which significantly prohibits standardized and uniform assays. Therefore, the database's primary objective was the deployment of a web-interface to find primer and probe information on experimentally validated qPCR assays submitted by colleagues in the field of real-time PCR (1).

RTPrimerDB provides unique identifiers (RTPrimerDB IDs) for validated qPCR assays preferentially published in peer-reviewed journals. The data that are maintained include all the information required to understand the purpose of an assay and to implement them in an experiment. These consist of gene and species nomenclature of the target sequence provided by Entrez Gene (2) and Ensembl (3), the primer and probe sequences (if any), the application and detection chemistry of the assay, the annealing temperature, details of the submitter and a wealth of links to databases for genes, protein domains, Single Nucleotide Polymorphisms (SNP) and publications (3–5). Additionally, for gene expression assays extra information is provided under the form of a graphical representation of the alignment of primers and probes on the known gene transcripts, supplemented with an amplicon secondary structure analysis module based on the mfold algorithm (6). Finally, users who tested an assay from the database can give their valuable feedback on assay performance in terms of specificity, sensitivity and reaction efficiency.

Currently, RTPrimerDB's functionality goes beyond its primary goal and provides an *in silico* analysis pipeline that performs streamlined quality assessment of newly designed qPCR primer pairs before their synthesis and experimental evaluation.

*To whom correspondence should be addressed. Tel: +32 9 2405187; Fax: +32 9 2406549; Email: Joke.Vandesompele@UGent.be

© The Author 2006. Published by Oxford University Press. All rights reserved.

The online version of this article has been published under an open access model. Users are entitled to use, reproduce, disseminate, or display the open access version of this article for non-commercial purposes provided that: the original authorship is properly and fully attributed; the Journal and Oxford University Press are attributed as the original place of publication with the correct citation details given; if an article is subsequently reproduced or disseminated not in its entirety but only in part or as a derivative work this must be clearly indicated. For commercial re-use, please contact journals.permissions@oxfordjournals.org

Downloaded from http://nar.oxfordjournals.org/ at Michigan State University on December 4, 2015

DATABASE STRUCTURE

Data retrieval

The primer and probe information can be accessed in two ways: a quick search uses a simple query on organism and gene symbol or RTPrimerDB ID, while more advanced searches can be performed based on gene name, official or alias gene symbol, detection chemistry, application, primer or probe sequence, PubMed ID or submitter's name. The search result page shows a list of the assays matching the search query with links to the individual assay reports or to the gene reports in which all assays for that particular gene are grouped together.

Gene reports

A gene report groups all available assays for a particular gene to help the evaluation of their usefulness in a specific experimental context (Figure 1). The overview starts with gene annotation data and links to Entrez Gene and Ensembl, followed by a graphical representation of the aligned primer and probe sequences onto the different transcript variants available from Ensembl. The occurrence of SNPs in the region where the primer/probe anneals and the exact location of the amplicon with respect to the gene exon structure or the protein domains can be examined at a glance, along with an assembly of other available assays for the same gene. Hovering over the SNP sites, the exon–intron structure, the protein domains and the primers and probes will display additional information on the bottom of the figure. The primer, probe and amplicon depiction contains a link to the individual assay report (see below). The known protein domains are linked to the Pfam database (4) and SNPs are linked to dbSNP (7). All figures are displayed as scalable vector graphics (SVG), a new web standard for zoomable and customizable high content vector graphics. SVG figures require installation of an appropriate browser plug-in.

The last part of the gene report gives an overview of all available assays for the selected gene and denotes if the primer pair is aligned to one or more of the known transcripts. The absence of an assay on the graphical mapping can be attributed to the specific alignment of one of the primers or probes in introns (e.g. for DNA quantification assays), untranscribed regions or alternative splice variants of the gene, or to sequence errors in the primers or probes.

Assay reports

The core of the database consists of assay reports with all information on primer and probe sequences that constitute one assay. Like the gene report, gene nomenclature information is provided together with a zoomed view on the alignment of the primers and probes on the various known gene transcripts (Figure 2).

The graphics contain the same functionality as described earlier. For each transcript, a link to the annotated sequence of the amplified region is available. Furthermore, a module to automatically evaluate the secondary DNA structure of a PCR amplicon is implemented, as this has been shown to be a critical factor for the efficiency of a PCR (6,8). The 'mfold' button directs to a page where secondary structure analysis is performed under default ionic conditions at the reported annealing temperature of the primers (Supplementary Figure 1). The occurrence of more or less stable hairpin structures around primer or probe annealing sites is automatically reported and visualized. Re-predicting the secondary structure formation can be done by adjusting the Mg^{++} or Na^{+} concentrations or annealing temperature.

Further, the primer and probe sequences are displayed together with the reported annealing temperature. The publication in which the assay is described in detail can be consulted by following the link to the PubMed literature database.

Last but not least, valuable feedback on assay performance from users who tested an assay from the database is displayed at the bottom of the page. Both the experimental evaluation details provided by the submitter as well as user's feedback will allow a better assessment of the reliability of an individual assay. The feedback includes details on the evaluation of the amplicon specificity based on melting curve analysis, gel electrophoresis, sequencing or restriction digest. Standard curve parameters (amplification efficiency, limit of detection, number of dilution points, dilution factor, correlation coefficient, etc.) are helpful to assess the efficiency and sensitivity of the particular assay.

DATA SUBMISSION

Users are invited to submit their validated assays after free registration, such that other users can benefit from their expertise. Upon successful login, an extra purple navigation bar appears on top of each page that guides the users to the advanced functionalities of RTPrimerDB. One of the links leads to an assay submission scheme suited for the addition of small numbers of primer pairs. Large dataset submissions should be performed by completing a special form located in the download section of the database.

In silico assay evaluation analysis pipeline

A second additional and recently developed functionality for registered users is an evaluation pipeline for newly designed primer and probe sets for gene expression analysis of human, mouse and rat genes. The '*in silico* evaluation' link provides a form to be completed with the target organism and gene, the primer and/or probe sequences and the annealing temperature of the designed primer pair. The next page gives detailed information of the alignment of the oligonucleotides onto the reference sequences of the different transcripts along with an indication of occurring SNP sites in the annealing regions (Supplementary Figure 2). There is also a direct link to NCBI's BLAST server with the settings applicable for short nearly exact sequence matches. Proceeding to the next page reveals the visual alignment of the primer pair on the transcript variants together with a link to the mfold analysis module already described.

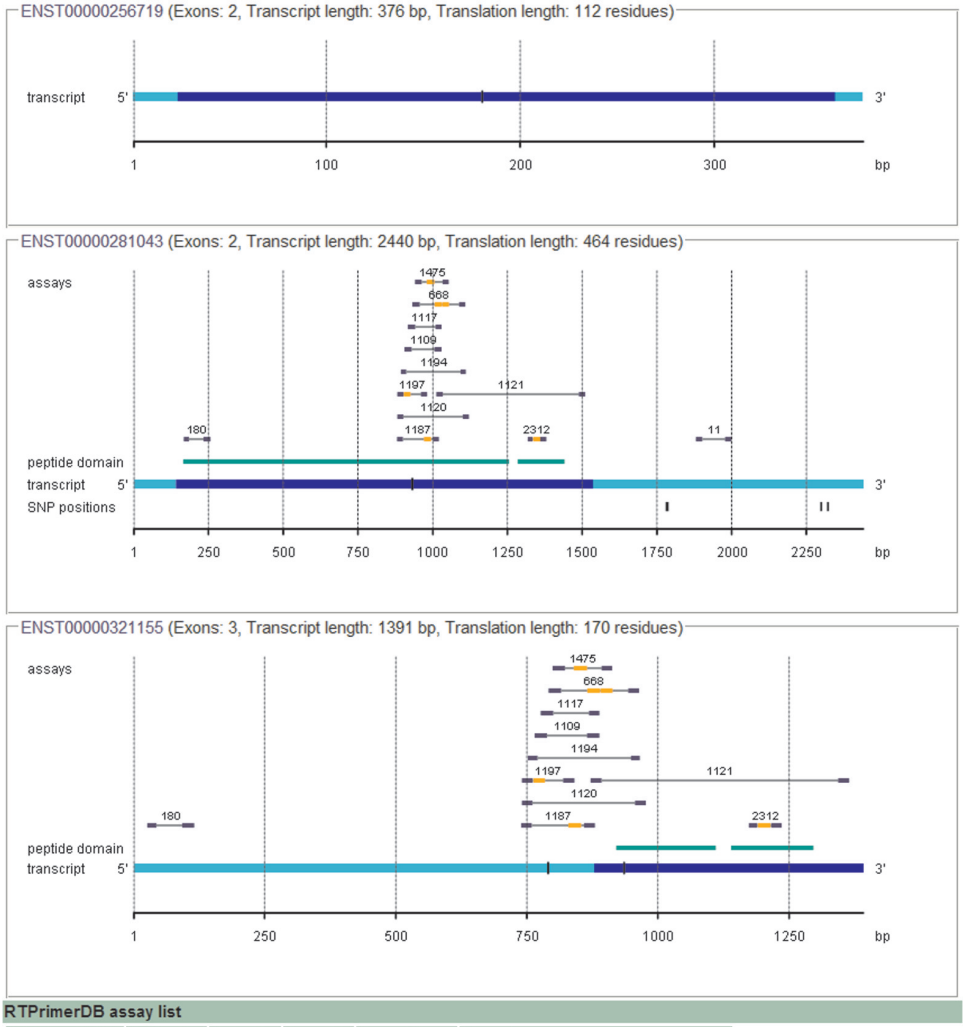
Executed evaluations are stored in the system for comparison of multiple potential primer sets to create a new assay (Supplementary Figure 3). All information in this assay evaluation analysis system is password protected and only viewed by the owner of the information.

Downloaded from http://nar.oxfordjournals.org/ at Michigan State University on December 4, 2015

Gene Information

MYCN: v-myc myelocytomatosis viral related oncogene, neuroblastoma derived (avian)
Alias Symbol(s): NMYC, N-myc
Organism: ***Homo sapiens*** (Hs, Human)
Entrez Gene ID: 4613
Ensembl Gene ID: ENSG00000134323

RTPrimerDB gene expression assay viewer
3 different transcripts are known for MYCN :



RTPrimerDB assay list

RTPrimerDB ID	Application	Detection	Template	Assay Viewer	Remarks
180	GXP	SYB	cDNA	YES	primers in coding sequence of MYCN
1109	GXP	SYB	cDNA	YES	
1117	GXP	SYB	cDNA	YES	
1120	GXP	SYB	cDNA	YES	
1187	GXP	TQ	cDNA	YES	
1194	GXP	SYB	cDNA	YES	
1475	GXP	TQ	cDNA	YES	
2312	GXP	TQ	cDNA	YES	
1197	GXP	TQ	cDNA	YES	amplicon located on exon 2 and 3
1121	GXP	SYB	cDNA	YES	
668	GCN	HYB	DNA	YES	
1195	GCN	TQ	DNA	NO	amplicon located in intron 2
11	GCN	SYB	DNA	YES	

Figure 1. A gene report consists of three major parts: (i) gene nomenclature information and links; (ii) the graphical mapping of the assays onto the different transcript variants and (iii) the listing of all available assays for that gene. Note that assay 1195 is not mapped on the transcripts because it's targeted to an intron.

Downloaded from http://nar.oxfordjournals.org/ at Michigan State University on December 4, 2015

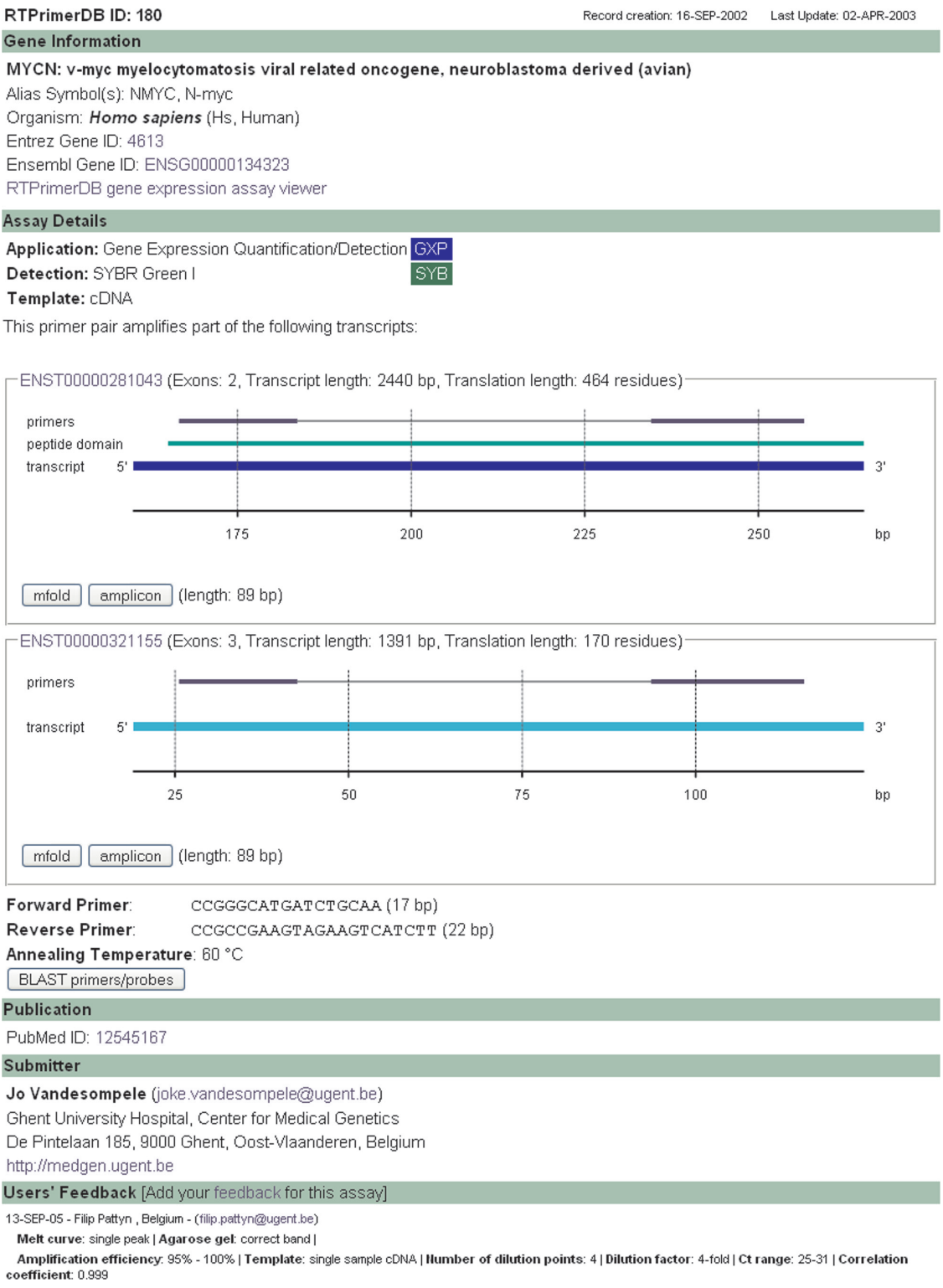


Figure 2. Assay report for RTPrimerDB ID 180 consists of five parts containing gene annotation data, graphical alignment of primer and probe sequences on the transcript variants, a publication reference, the submitter's contact details and experimental feedback provided by users.

SOFTWARE REQUIREMENTS

RTPrimerDB is fully supported by all recently updated versions of internet browsers with a graphical interface on Windows, Macintosh, Linux and Unix operating systems. The SVG graphics are visible after installing a freely available version of an SVG viewer. More information is provided in the FAQ section of RTPrimerDB.

Downloaded from http://nar.oxfordjournals.org/ at Michigan State University on December 4, 2015

FUTURE DIRECTIONS

RTPrimerDB will continually be updated with new detection chemistries and applications that become generally established. Immediate plans include a major extension of the number of species selectable by registered users to submit new primer assays. Also the *in silico* evaluation analysis system will be extended to analyse primers against all species currently available in Ensembl. The mapping of primers and probes for genomic assays in an assay viewer will be made available in a new release of the database.

FEEDBACK

We welcome your feedback with respect to the RTPrimerDB interface, or any data contained therein. You may use the feedback form available from each page or send comments to RTPrimerDB@medgen.UGent.be.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We greatly acknowledge the help from Jelle Verspurten and insightful suggestions from Stephen Bustin. F.P. is a Research Assistant and J.V. a Postdoctoral Researcher of the Research Foundation—Flanders (FWO—Vlaanderen). This study is supported by GOA-grant 12051203, FWO-grant G.0185.04, G.1.5.243.05 and G.0106.05, and research grant of 'Kinderkankerfonds' (a non-profit childhood cancer foundation under Belgian law). This text presents research results of the Belgian program of Interuniversity Poles of Attraction initiated by the Belgian State, Prime Minister's Office, Science Policy Programming (IUAP). Funding to pay the Open Access publication charges for this article was provided by Ghent University BOF-FWO bench fee.

Conflict of interest statement. None declared.

REFERENCES

1. Pattyn,F., Speleman,F., De Paepe,A. and Vandesompele,J. (2003) RTPrimerDB: the real-time PCR primer and probe database. *Nucleic Acids Res.*, **31**, 122–123.
2. Maglott,D., Ostell,J., Pruitt,K.D. and Tatusova,T. (2005) Entrez Gene: gene-centered information at NCBI. *Nucleic Acids Res.*, **33**, D54–D58.
3. Hubbard,T., Andrews,D., Caccamo,M., Cameron,G., Chen,Y., Clamp,M., Clarke,L., Coates,G., Cox,T., Cunningham,F. *et al*. (2005) Ensembl 2005. *Nucleic Acids Res.*, **33**, D447–D453.
4. Bateman,A., Coin,L., Durbin,R., Finn,R.D., Hollich,V., Griffiths-Jones,S., Khanna,A., Marshall,M., Moxon,S., Sonnhammer,E.L. *et al*. (2004) The Pfam protein families database. *Nucleic Acids Res.*, **32**, D138–D141.
5. Wheeler,D.L., Barrett,T., Benson,D.A., Bryant,S.H., Canese,K., Church,D.M., DiCuccio,M., Edgar,R., Federhen,S., Helmberg,W. *et al*. (2005) Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res.*, **33**, D39–D45.
6. Zuker,M. (2003) Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.*, **31**, 3406–3415.
7. Sherry,S.T., Ward,M.H., Kholodov,M., Baker,J., Phan,L., Smigielski,E.M. and Sirotkin,K. (2001) dbSNP: the NCBI database of genetic variation. *Nucleic Acids Res.*, **29**, 308–311.
8. Hoebeeck,J., van der Luijt,R., Poppe,B., De Smet,E., Yigit,N., Claes,K., Zewald,R., de Jong,G.J., De Paepe,A., Speleman,F. *et al*. (2005) Rapid detection of VHL exon deletions using real-time quantitative PCR. *Lab. Invest.*, **85**, 24–33.

Downloaded from http://nar.oxfordjournals.org/ at Michigan State University on December 4, 2015